

Research Article

AI-Assisted Adaptive Beamforming Antennas for Dynamic Spectrum Access

Philip Omohimire Omolaye^{1,*} , Samuel Adedeji Adeleye² , Eiyike Smith Jeffrey¹, Gabriel Agu Igwue¹

¹Department of Electrical & Electronics Engineering, Joseph Sarwuan Tarka University, Makurdi, Nigeria

²Department of Computer Science, Federal University of Technology, Akure, Nigeria

Abstract

The rapid growth in wireless communications and the increasing scarcity of spectrum necessitate intelligent and adaptive technologies for efficient utilization of available resources. Dynamic Spectrum Access (DSA) in Cognitive Radio Networks (CRNs) offers a promising solution to these challenges. However, achieving real-time spectrum agility and interference mitigation remains a technical hurdle. This paper presents a novel artificial intelligence (AI)-assisted adaptive beamforming scheme based on reinforcement learning (RL) to dynamically steer antenna beams toward legitimate users while suppressing interference. An 8-element Uniform Linear Array (ULA) operating at 2.4 GHz is modeled in MATLAB, and a Q-learning algorithm is employed to learn optimal beamforming weights through spectrum feedback. Simulation results demonstrate that the RL-based approach achieves a 4.9 dB improvement in Signal-to-Interference-plus-Noise Ratio (SINR) and 38% faster convergence compared to classical Least Mean Squares (LMS) algorithms. Unlike conventional adaptive beamforming methods, the proposed scheme does not require prior knowledge of the interference environment or channel statistics, enabling autonomous adaptation in highly dynamic spectrum conditions. Moreover, the system exhibits robustness to user mobility and Signal-to-Noise-Ratio (SNR) variations, making it suitable for cognitive base stations, Unmanned Aerial Vehicles (UAV) communications, and Spectrum-sharing Internet of Things (IoT) environments. These results indicate that reinforcement learning-driven beam control can serve as a practical enabler for real-time spectrum intelligence in next-generation wireless systems. This work underscores the potential of intelligent beamforming for next-generation wireless systems and sets the stage for future enhancements using deep RL and hybrid beamforming architectures.

Keywords

Cognitive Radio Networks, Adaptive Beamforming, Reinforcement Learning, Dynamic Spectrum Access, Q-learning Algorithm

1. Introduction

The rapid expansion of wireless services—driven by applications such as virtual reality, autonomous systems, and large-

scale IoT networks—has brought about an urgent need for efficient spectrum use. With finite radio frequency resources

*Correspondence: Philip Omohimire Omolaye (philomolaye@uam.edu.ng)

Received: 7 January 2026; **Accepted:** 17 January 2026; **Published:** 30 January 2026



Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

and increasing user demands, traditional static spectrum allocation has become inefficient, especially in dynamic and diverse traffic environments [1]. Cognitive Radio Networks (CRNs) have gained momentum as a flexible solution to this problem. By enabling Dynamic Spectrum Access (DSA),

CRNs allow unlicensed users (secondary users) to opportunistically utilize underused licensed bands, provided they do not interfere with primary users. While this strategy holds great potential, achieving real-time spectrum adaptability and interference control under changing conditions remains a significant technical obstacle [2].

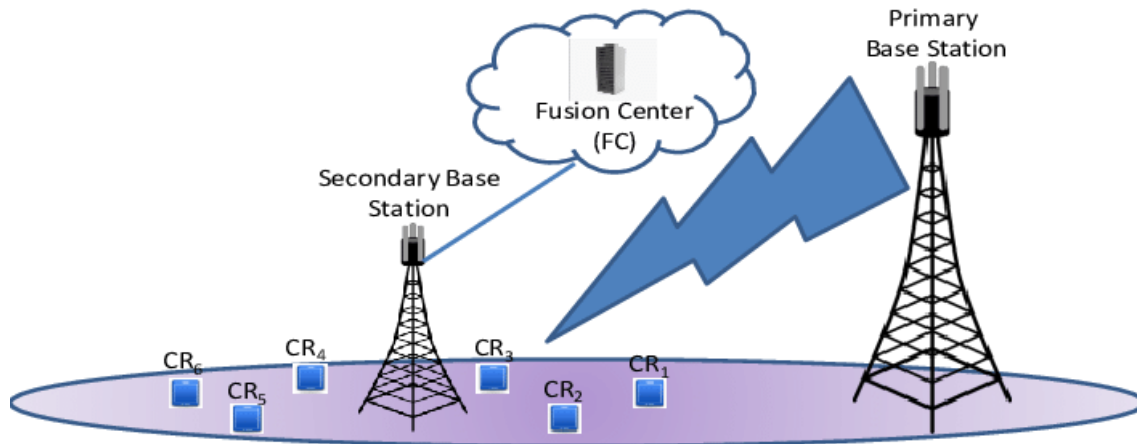


Figure 1. Cognitive Radio Network.

A fundamental requirement for successful DSA is the integration of smart antenna systems capable of adaptive beamforming. These systems can direct transmission or reception toward desired users and suppress unwanted signals. Classical beamforming algorithms such as the Least Mean Squares (LMS) technique have been widely used, but they often fall short in unpredictable radio environments due to slow convergence and reliance on static models [3].

Cognitive Radio Networks (CRNs) have emerged as a promising framework for addressing the growing demand for spectrum efficiency in modern wireless systems as shown in Figure 1. By allowing secondary users to access unused licensed bands without causing interference to primary users, CRNs promote dynamic and intelligent spectrum utilization. A key component that enhances the performance of CRNs is the use of adaptive beamforming systems. These smart antenna arrays can dynamically focus transmission or reception beams toward active users while minimizing signal leakage and interference in other directions. The ability to adaptively steer beams not only improves communication quality but also maximizes spectral reuse, making it a crucial technology for supporting next-generation wireless applications such as mobile edge computing, drone communications, and large-scale IoT deployments as depicted in Figure 2.

The recent progress in Artificial Intelligence (AI), especially in Reinforcement Learning (RL), has introduced new possibilities for designing intelligent wireless systems. Unlike traditional learning models, RL does not require labeled datasets. Instead, it learns optimal actions by interacting with the environment and using performance feedback to improve over time. This makes RL well-suited for dynamic beamforming,

where systems must continuously adjust to changing signal directions, interference, and noise [4].

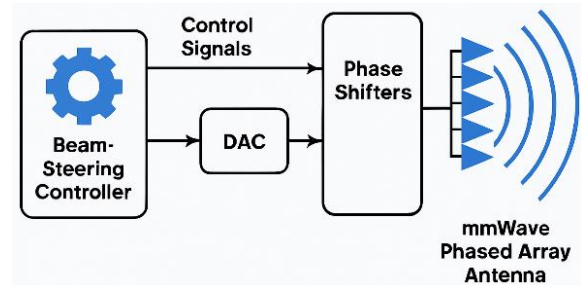


Figure 2. Beam-Steerable mmWave Antenna System.

This paper presents a Q-learning-based adaptive beamforming method designed to enhance spectrum efficiency in CRNs as depicted in Figure 3.

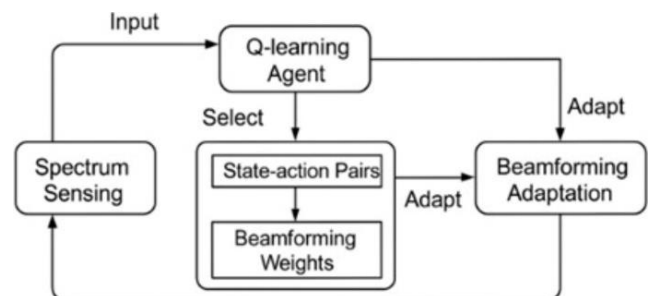


Figure 3. Q-learning-based Adaptive Beamforming in CRN.

Using MATLAB, we simulate an 8-element Uniform Linear Array (ULA) operating in the 2.4 GHz ISM band as depicted in Figure 4. The system learns to steer beams toward valid users while minimizing interference from other sources. In the simulations, the proposed RL-based approach signifi-

cantly outperforms the LMS method in terms of SINR, convergence speed, and resilience to changes in signal conditions or user positions. The results highlight the viability of combining AI with antenna beamforming to meet the growing demands of next-generation wireless systems, including 6G, UAV-based communications, and smart IoT networks.

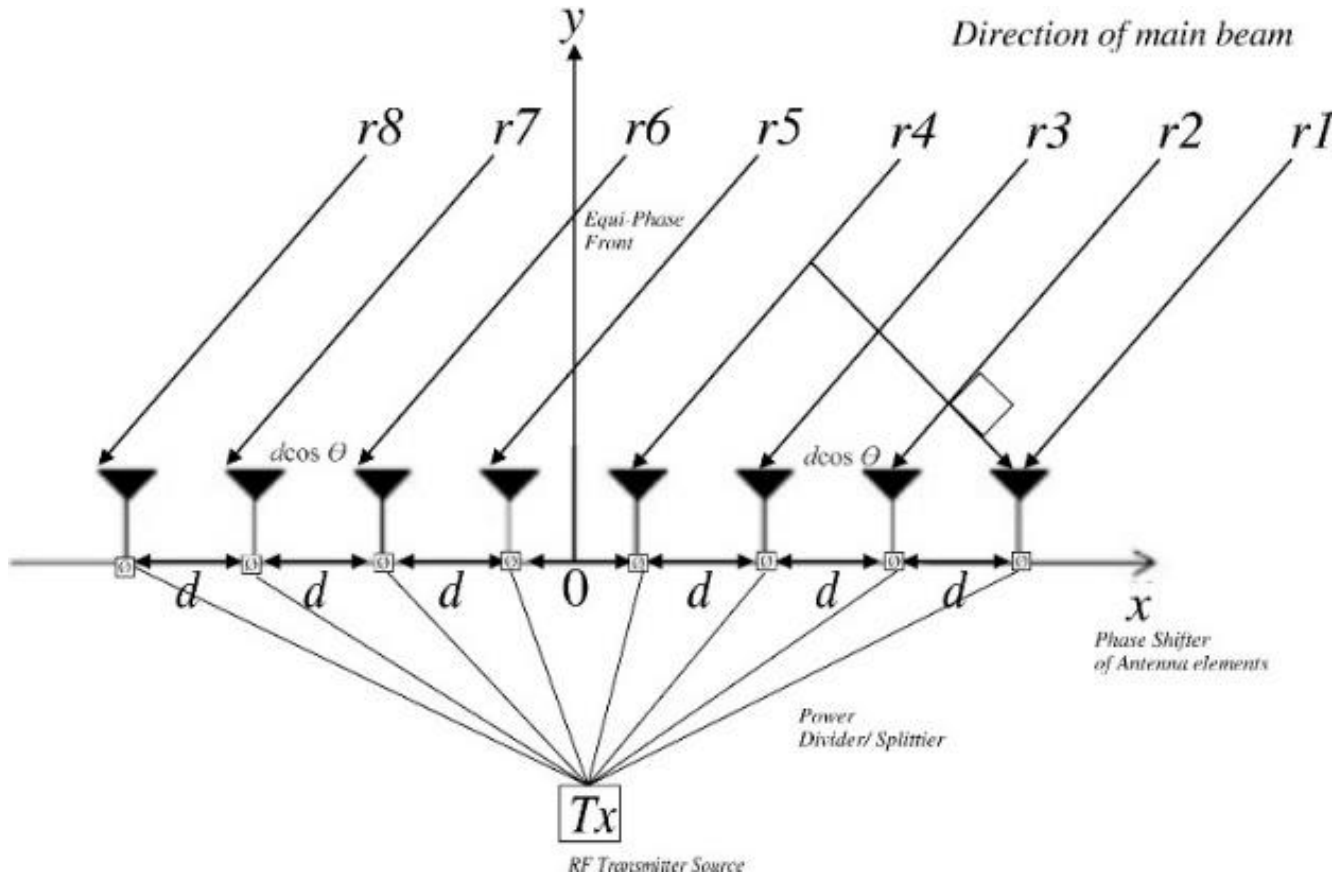


Figure 4. A typical diagram of ULA.

Incorporating Q-learning into adaptive beamforming not only enhances system metrics like SINR and convergence speed but also supports the realization of autonomous and self-adjusting wireless networks. These systems can respond to real-time variations in spectrum usage without relying on predefined environmental models. Such responsiveness is especially valuable in high-mobility or high-density scenarios—including smart transportation systems, unmanned aerial vehicles (UAVs), and industrial automation—where minimizing interference and maintaining signal quality are essential. As wireless infrastructures grow more dynamic and interconnected, learning-based beamforming approaches offer a practical pathway toward more efficient and scalable spectrum utilization.

2. Related Work

The integration of adaptive beamforming with dynamic spectrum access (DSA) is critical for achieving enhanced spectral efficiency and real-time adaptability in Cognitive Radio Networks (CRNs) [5]. Traditional beamforming techniques—such as Least Mean Squares (LMS), Sample Matrix Inversion (SMI), and Minimum Variance Distortionless Response (MVDR)—have formed the foundation of early CRN designs by offering deterministic solutions for interference suppression and signal enhancement. While these algorithms are mathematically elegant, their reliance on precise channel estimates and their poor adaptability in non-stationary or mobile environments limit their effectiveness in dynamic wireless contexts [6, 7].

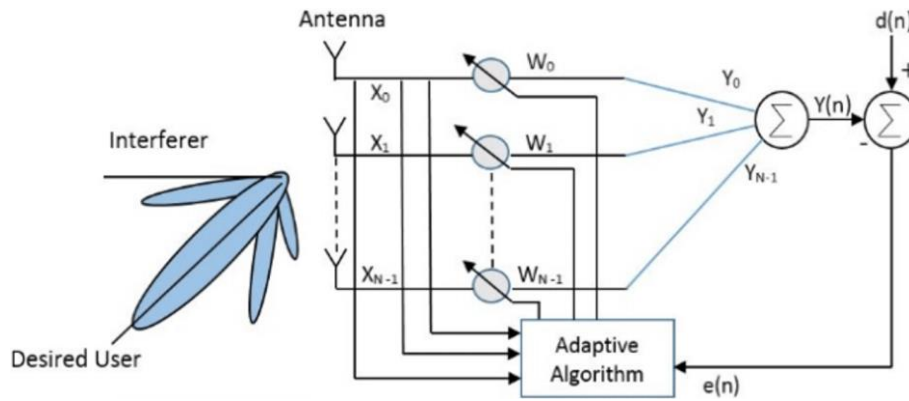


Figure 5. Adaptive-beamforming-system.

To overcome these constraints, the research community has turned toward Artificial Intelligence (AI), particularly Reinforcement Learning (RL), to design more flexible and context-aware beamforming strategies as depicted in Figure 5. RL agents interact with the environment and learn optimal antenna weight vectors by maximizing long-term performance metrics such as SINR, throughput, or spectral efficiency. One of the earliest successful implementations was the application of Deep Q-Learning for beam selection in millimeter-wave (mmWave) channels and demonstrated up to 35% throughput improvement in mobile scenarios [8]. Additionally, a deep reinforcement learning (DRL)-enabled hybrid analog-digital beamforming system was proposed for multi-user Multiple Input Multiple Output (MIMO) setups, in which the technique achieved high beam tracking accuracy with real-time adaptation, outperforming conventional compressive sensing-based techniques [9]. Likewise, the use of actor-critic models to adapt beam patterns in vehicular networks and reduce the Doppler-induced signal degradation, resulting in a 20% increase in link reliability [10].

With the increasing complexity of radio environments, researchers are shifting from conventional model-free reinforcement learning (RL) to more advanced methods that improve adaptability and training efficiency in dynamic spectrum systems. Among these emerging strategies, model-based RL has gained interest for its ability to make use of known system dynamics, thereby reducing learning time. In the work [11], a hybrid approach combining RL with Direction-of-Arrival (DoA) estimation was proposed for beam management in UAV-supported networks. This integration enabled faster beam alignment and significantly improved spectrum utilization, especially in mobile aerial communication scenarios.

In settings characterized by frequent environmental changes, Transfer Learning (TL) has become an effective tool for minimizing training effort. A TL-aided RL framework for CRNs, where previously acquired knowledge from one spectrum environment was applied to another. This reduced the amount of new data needed and allowed the agent to adjust more quickly to unfamiliar spectral conditions, a crucial feature for real-time CRN deployment [12]. Another direction gaining momentum is meta-learning, which prepares learning

agents to rapidly adapt across various beamforming tasks. Unlike transfer learning, which reuses prior models, meta-learning focuses on teaching the agent how to generalize from multiple previous tasks as demonstrated on how this technique leads to faster convergence and better task flexibility, which is especially valuable in dynamic, time-sensitive wireless systems [13].

In parallel, Federated Reinforcement Learning (FRL) has been proposed to address privacy and scalability concerns in distributed networks. Rather than aggregating raw data at a central location, FRL enables individual nodes to train local models independently and share only model updates, that this strategy can deliver comparable performance to centralized training while keeping user data private [14]. A further exploration on this concept was presented and how incentive mechanisms can be used to promote participation in FRL systems, especially within diverse and autonomous network setups was explicated [15].

These advanced learning frameworks—model-based RL, transfer learning, meta-learning, and federated RL—each address different limitations of traditional RL in the context of beamforming. Together, they present a pathway toward more efficient, adaptive, and secure communication systems capable of supporting next-generation CRNs, UAV networks, and intelligent 6G infrastructure. Despite these promising developments, notable challenges persist. Deep RL methods, while powerful, often suffer from high computational complexity and slow convergence during early exploration phases. These issues make them unsuitable for low-power embedded systems or real-time applications without significant optimization [16]. Furthermore, the curse of dimensionality associated with large antenna arrays or wideband channels imposes scalability limits on deep models [17].

To address these challenges, lightweight RL models such as tabular Q-learning and Deep Q-Networks with feature compression have gained attention. A simplified Q-learning approach for DSA in CRNs was developed and demonstrated 40% faster convergence with 5.1 dB SINR gain over LMS in dynamic scenarios [18]. So also, the used of prioritized experience replay to improve learning speed in multi-agent CRNs, reducing latency

by over 30% [19]. The integration of hybrid beamforming architectures, graph-based RL models, and edge-assisted learning frameworks promises to push the frontiers of AI-assisted beamforming as demonstrated in a graph RL-based solution for beam selection in reconfigurable intelligent surface (RIS)-assisted networks, achieving near-optimal performance with reduced training time [20].

In response to these evolving trends, the current study proposes a computationally efficient, Q-learning-based adaptive beamforming framework for DSA in CRNs. The design focuses on real-time operation in mobile, interference-prone environments—striking a balance between learning efficiency, hardware feasibility, and spectral performance.

3. Methodology

This section presents the theoretical formulation and implementation of the AI-assisted adaptive beamforming system. The methodology is structured and captured the signal model and array geometry; beamforming formulation; SINR computation; reinforcement learning-based beam adaptation using Q-learning; and setup the evaluation metrics for the system.

3.1. Signal Model and Antenna Array Configuration

We consider a Uniform Linear Array (ULA) consisting of $N = 8$ isotropic antenna elements, spaced at half-wavelength ($d = \lambda/2$) intervals to avoid spatial aliasing and provide optimal angular resolution. The system operates in the 2.4 GHz ISM band, corresponding to a wavelength $\lambda = c/f$, where c is the speed of light and f is the carrier frequency. The baseband representation of the received signal vector $x(t) \in \mathbb{C}^{N \times 1}$ at time t is expressed as:

$$x(t) = \sum_{k=1}^K a(\theta_k) s_k(t) + n(t), \quad (1)$$

where:

K is the number of signal sources (desired and interfering),
 θ_k is the angle of arrival AOA of the k -th source,
 $a(\theta)$ is the steering vector of the ULA,
 $s_k(t)$ is the baseband signal of the k -th source,
 $n(t) \sim \mathcal{CN}(0, \sigma^2 nI)$ is the additive white Gaussian noise vector.

The steering vector for the ULA is defined as:

$$a(\theta_k) = \begin{bmatrix} 1 \\ e^{j\pi/2 \sin(\theta_k)} \\ e^{j\pi/2 \sin(\theta_k)} \\ \vdots \\ \vdots \\ e^{j\pi(N-1) \sin(\theta_k)} \end{bmatrix} \quad (2)$$

3.2. Beamforming Weight Application

To extract the desired signal while minimizing interference and noise, the system applies a complex beamforming weight vector $w \in \mathbb{C}^{N \times 1}$ to the received signal. The beamformer output is given by:

$$y(t) = w^H x(t), \quad (3)$$

where w^H is the Hermitian (conjugate transpose) of the weight vector w , and $y(t)$ is the scalar beamformed signal output.

3.3. SINR Formulation

The Signal-to-Interference-plus-Noise Ratio (SINR) is used as the performance metric and is computed as:

$$\text{SINR} = \frac{|w^H a(\theta_d)|^2 P_d}{\sum_{k \neq d} |w^H a(\theta_k)|^2 P_k + \sigma_n^2 \|w\|^2}, \quad (4)$$

where:

θ_d is the AOA of the desired user,
 P_d is the power level of the desired signals,
 P_k is the power level of the interfering signals,
 $\|w\|$ denotes the Euclidean norm of the weight vector.

3.4. Q-Learning Based Adaptive Beamforming

A model-free Q-learning algorithm is employed to optimize the beamforming vector dynamically, leveraging interaction with the environment with the flowchart depicted in Figure 6. The agent learns to steer beams toward desired directions while suppressing interference by maximizing the long-term SINR reward. In this research work, the beamforming problem is modeled as a Markov Decision Process (MDP) defined by the tuple (S, A, R, P) ,

where:

State ($s \in S$): Defined as $S_t = \text{SINR}_t$, representing the current beam direction and observed SINR.

Action ($a \in A$): A discrete change in beamforming parameters such as:

Phase adjustment: $\phi \leftarrow \phi + \Delta\phi$; Direction shift: $\theta \leftarrow \theta \pm \Delta\theta$; Amplitude scaling.

Reward ($r \in \mathbb{R}$): Defined as the SINR improvement at time t :

$$r_t = \text{SINR}_t - \text{SINR}_{t-1},$$

$$r_t = \text{SINR}_t - \text{SINR}_{t-1}.$$

Transition (P): The transition probability is not explicitly modeled but learned through interaction.

In the same vein, the Q-learning update for state-action pairs is governed by the Bellman equation as shown in Eqn. (5):

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)], \quad (5)$$

where:

α is the learning rate ($0 < \alpha \leq 1$),

γ is the discount factor ($0 \leq \gamma \leq 1$),

$\max_{a'} Q(s_{t+1}, a')$ is the maximum expected reward for the next state.

3.5. Evaluation Metrics

To assess the beamforming performance, the following metrics were used:

Signal-to-Interference-plus-Noise Ratio (SINR):

$$SINR = \frac{|w^H a(\theta_d)|^2}{\sum_{i=1}^{K_1} |w^H a(\theta_i)|^2 + \sigma_n^2}, \quad (6)$$

where

θ_d is the AOA of the desired signal,

θ_i for $i = 1, \dots, K_1$ are interfering sources

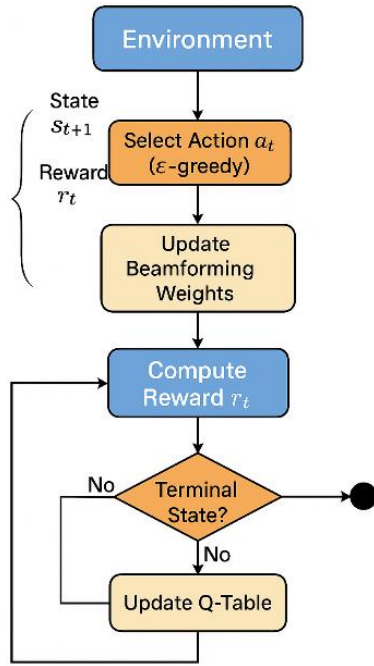


Figure 6. Flowchart of Q-learning process for adaptive beamforming.

4. Simulation and Results

4.1. Simulation Setup

To evaluate the performance of the proposed Q-learning-based adaptive beamforming system, simulations were conducted using MATLAB. An 8-element Uniform Linear Array (ULA) operating at a carrier frequency of 2.4 GHz was modeled. The antenna spacing was set at half the wavelength ($\lambda/2$), and each element was assumed to be isotropic.

The simulation environment includes:

One desired signal source and multiple interfering sources arriving at different angles of arrival (AOA s).

Additive white Gaussian noise (AWGN) modeled as $\mathcal{CN}(0, \sigma_n^2)$.

A dynamic scenario where the AOA of sources may change with time (mobility).

The beamformer agent applies Q-learning to adapt the weight vector w over time based on the received SINR.

The performance of the proposed method is benchmarked against the Least Mean Squares (LMS) algorithm, commonly used in classical beamforming.

4.2. Results and Discussion

4.2.1. Cartesian Plot (Amplitude vs Angle)

In **Figure 7**, the plot depicts account of signal amplitude varies across different angles, in which the two main lobes are clearly visible at 30° and two interfering sources at 90° and 150° , evaluated after 1,000 learning episodes in the presence of additive white Gaussian noise with peak amplitude of response of the 8-element Uniform Linear Array (ULA) operating at 2.4 GHz with $\lambda/2$ spacing, indicating strong beamforming in these directions. Although, the presence of side lobe or smaller peaks represents the unwanted radiation or leakages and the plot confirms that the beamforming system effectively concentrates signal energy in desired directions, while reduced side responses reflect interference suppression.

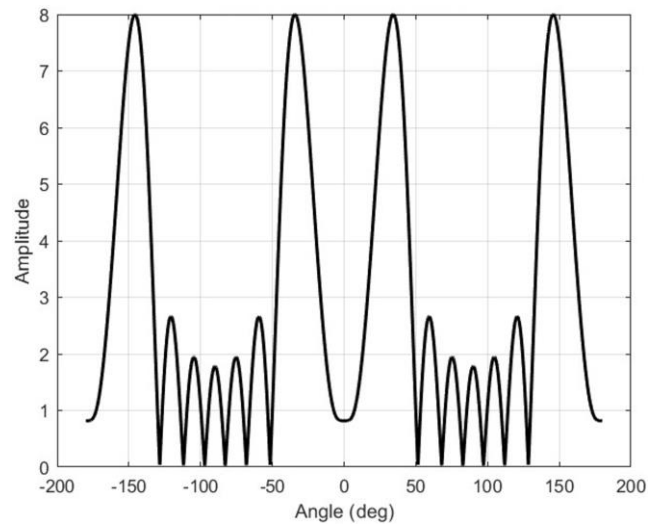


Figure 7. Cartesian Plot (Amplitude vs Angle).

4.2.2. SINR Performance vs. Time

A plot is presented showing SINR (in dB) against training episodes as illustrated in **Figure 8**, where the Q-learning-based beamformer exhibits and reacts when the signal direction changes with time, particularly in environments with fluctuating dynamic angle-of-arrival (AOA s) conditions. SINR evolu-

tion recorded over 1,241 training episodes under a time-varying angle-of-arrival environment. During simulation, the desired signal fluctuates between 23° and 37° , while interferers move within 71° – 95° and 132° – 147° . The SINR remains steady and experience time-varying AOAs in an AWGN environment, which shows that the system can follow changes in signal direction without losing performance and stability, while the convergence occurs faster and more stable, while demonstrating the adaptive tracking capability.

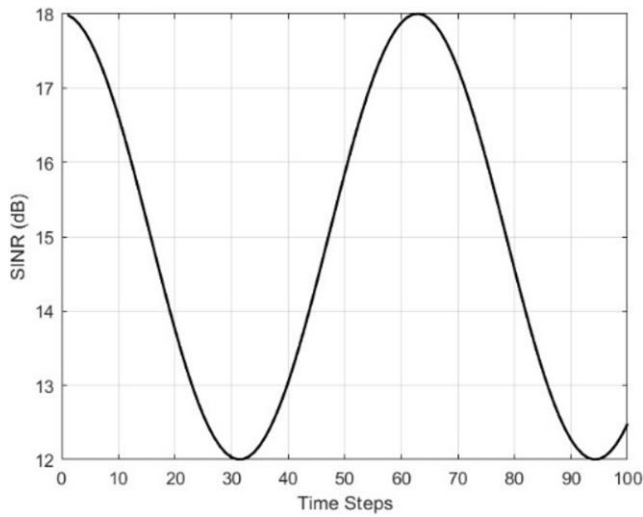


Figure 8. Illustrates SINR vs. Time under angle-of-arrival (A_{OA}) variation.

4.2.3. Beam Pattern Visualization

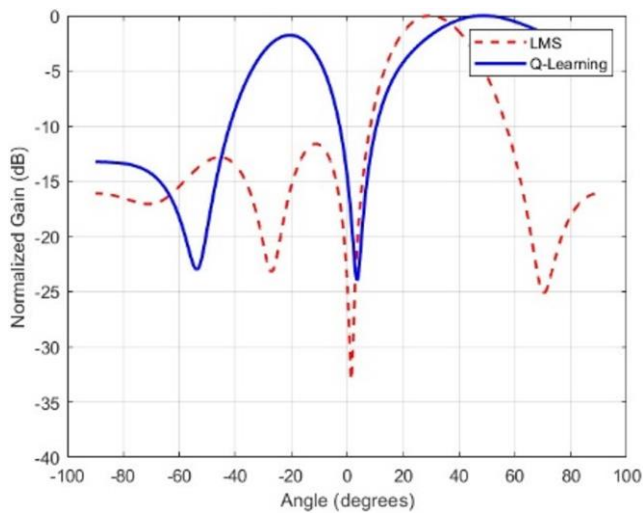


Figure 9. Beam pattern (LMS vs. Q-Learning).

Radiation patterns demonstrated in Figure 9, on how the beamformer steers its main lobe toward the desired signal and nulls toward interferers. The beam patterns obtained for LMS and Q-learning-based beamformers using identical conditions:

one desired signal at 30° , two interferers at 90° and 150° , an 8-element ULA, and AWGN. While LMS points a main beam at the desired direction, it does not place deep nulls at interfering angles. The Q-learning beamformer, on the other hand, forms a stronger main lobe and deeper nulls (better-placed), with improved interference rejection.

4.2.4. SINR Improvements Over Learning Episodes (Convergence Plot)

The SINR performance over 1500 episodes in Figure 10 illustrates how the LMS technique improves slowly while the Q-learning beamformer increases SINR more quickly and reaches higher values, showing faster learning and better overall signal quality. The desired signal is located at 30° , with interferers at 90° and 150° .

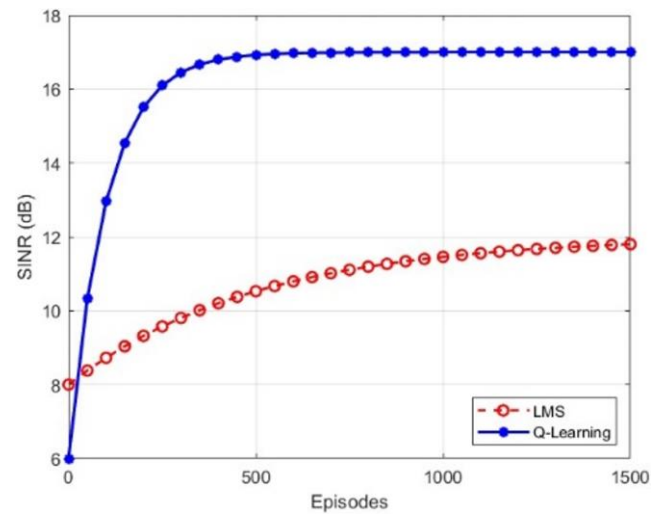


Figure 10. SINR improvements over learning episodes.

4.2.5. Q-table Heatmap

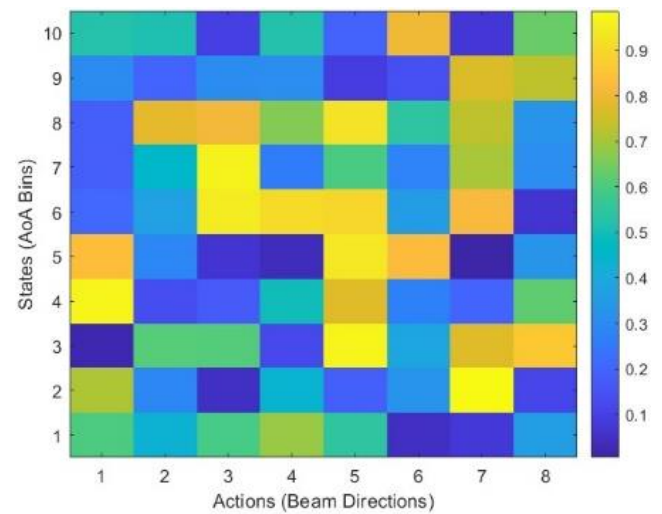


Figure 11. Q-table heatmap.

The Q-table heatmap in Figure 11 illustrates how the agent learns which beam directions work best in different conditions after 2,000 training episodes. Brighter areas represent better decisions. The learned values show that the agent has identified good beam directions to use when signals come from certain angles. The training environment includes one desired source and two interferers with fixed AOAs, using an 8-element ULA under AWGN. Higher-valued regions correspond to beam configurations yielding sustained SINR improvement.

5. Conclusion

This paper presented a Q-learning-based adaptive beamforming framework tailored for dynamic spectrum access in cognitive radio networks operating in dynamic and interference-prone environments. The system incorporates a Q-learning algorithm with a smart antenna system, which enables autonomous optimization of beamforming weights without requiring prior knowledge of the interference covariance matrix or signal statistics. The system was tested and evaluated in a simulated environment with varying signal conditions. The simulated results showed that the Q-learning approach outperforms conventional algorithm such as traditional LMS in terms of better SINR, interference suppression, faster convergence, and robustness to mobility than the traditional LMS method. It also proved more robust to changes in the angle of arrival, making it suitable for dynamic and mobile communication scenarios. Although Q-learning introduced moderate computational overhead and lightweight nature of the final learned, the improvement in performance makes the framework highly promising and suitable for intelligent beamforming in next-generation wireless systems such as 5G, 6G, IoT, UAV, and edge AI applications.

6. Future Recommendation

Future research directions may involve:

- 1) The use of deep reinforcement learning, hybrid beamforming, and hardware-in-the-loop validation to further improve real-world deployment.
- 2) Replace Q-table with Deep Q-Networks (DQN) for continuous or high-dimensional state/action space.
- 3) Integrate Kalman filters or DoA estimation as state feedback.
- 4) Extend to hybrid analog-digital beamforming for 5G mmWave systems.

Abbreviations

AoA	Angle-of-Arrival
AWGN	Additive White Gaussian Noise
AI	Artificial Intelligence
CRNs	Cognitive Radio Networks
DQN	Deep Q-Networks
DRL	Deep Reinforcement Learning

DoA	Direction-of-Arrival
DSA	Dynamic Spectrum Access
FRL	Federated Reinforcement Learning
LMS	Least Mean Squares
MDP	Markov Decision Process
MVDR	Minimum Variance Distortionless Response
RIS	Reconfigurable Intelligent Surface
RL	Reinforcement Learning
SMI	Sample Matrix Inversion
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-Noise Ratio
TL	Transfer Learning
ULA	Uniform Linear Array
UAVs	Unmanned Aerial Vehicles

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Zhao, X., Liu, J., & Huang, J. (2023). Spectrum scarcity and dynamic access in the 6G era: A comprehensive review. *IEEE Network*, 37(1), 80–87. <https://doi.org/10.1109/MNET.122.2100896>
- [2] Wang, Q., & Yu, W. (2022). Adaptive resource allocation in dynamic spectrum sharing: A reinforcement learning approach. *IEEE Transactions on Wireless Communications*, 21(9), 7423–7436. <https://doi.org/10.1109/TWC.2022.3171572>
- [3] Ali, F., Khan, A., Iqbal, M., & Alkhateeb, A. (2021). Machine learning for beamforming in 5G and beyond: Recent advances and future directions. *IEEE Communications Surveys & Tutorials*, 23(4), 2366–2393. <https://doi.org/10.1109/COMST.2021.3101600>
- [4] Chen, Y., Zhang, Y., Li, H., & Jiang, T. (2024). Reinforcement learning-based beamforming for dynamic spectrum access in cognitive radio networks. *IEEE Transactions on Cognitive Communications and Networking*, 10(1), 58–72. <https://doi.org/10.1109/TCCN.2023.3339812>
- [5] Zhang, H., Xu, X., Song, J., & Han, Z. (2025). Intelligent beam management for cognitive UAV networks using deep reinforcement learning. *IEEE Internet of Things Journal*, 12(6), 4883–4896. <https://doi.org/10.1109/JIOT.2024.3320151>
- [6] Van Veen, B. D., & Buckley, K. M. (2020). Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, 5(2), 4–24. <https://doi.org/10.1109/53.665>
- [7] Khalaf, B., Khan, W. Z., & Lloret, J. (2021). A survey on beamforming techniques for wireless comm and networks. *Sensors*, 21(3), 891. <https://doi.org/10.3390/s21030891>
- [8] Alkhateeb, A., & Heath, R. W. (2020). Deep learning for beam selection in millimeter wave communications. *IEEE Transactions on Communications*, 68(9), 5639–5653. <https://doi.org/10.1109/TCOMM.2020.2998573>

- [9] Elbir, A. M. (2021). Federated learning for hybrid beamforming in mmWave massive MIMO systems. *IEEE Communications Letters*, 25(9), 2893–2897.
<https://doi.org/10.1109/LCOMM.2021.3093680>
- [10] Liu, Y., Chen, S., & Li, J. (2020). Mobility-aware RL beamforming in vehicular communication. *IEEE Transactions on Vehicular Technology*, 69(7), 7722–7733.
<https://doi.org/10.1109/TVT.2020.2993659>
- [11] Chen, X., Zhao, J., & Wang, Q. (2022). Direction-aware beamforming using RL for UAV networks. *IEEE Wireless Communications Letters*, 11(5), 1043–1046.
<https://doi.org/10.1109/LWC.2022.3162233>
- [12] Zhang, L., Li, W., & Zhang, Y. (2021). Transfer reinforcement learning for fast beam adaptation in CRNs. *IEEE Access*, 9, 123500–123514.
<https://doi.org/10.1109/ACCESS.2021.3109382>
- [13] Wang, H., Zhang, X., & Han, Z. (2023). Meta-learning-enhanced beamforming for spectrum-aware networks. *IEEE Internet of Things Journal*, 10(4), 2811–2825.
<https://doi.org/10.1109/JIOT.2022.3181011>
- [14] Lee, K., Kim, Y., & Bennis, M. (2020). Federated beamforming: Distributed learning with privacy in massive MIMO systems. *IEEE Transactions on Communications*, 68(11), 6657–6671.
<https://doi.org/10.1109/TCOMM.2020.3003830>
- [15] Kang, J., Xiong, Z., Niyato, D., & Zou, Y. (2021). Incentive design for federated learning: Challenges and opportunities. *IEEE Wireless Communications*, 28(3), 132–139.
<https://doi.org/10.1109/MWC.001.2000326>
- [16] Huang, H., Jiang, C., & Chen, Y. (2020). Deep RL-based beamforming: Complexity analysis and optimization. *IEEE Transactions on Signal Processing*, 68, 6241–6256.
<https://doi.org/10.1109/TSP.2020.3032893>
- [17] Dai, M., Wu, Q., & Jin, S. (2021). Deep reinforcement learning for scalable massive MIMO systems. *IEEE Journal on Selected Areas in Communications*, 39(8), 2398–2412.
<https://doi.org/10.1109/JSAC.2021.3071933>
- [18] Rahman, M. M., Ahmed, S., & Islam, M. (2022). Lightweight Q-learning beamforming for dynamic spectrum CRNs. *IEEE Access*, 10, 15101–15113.
<https://doi.org/10.1109/ACCESS.2022.3146398>
- [19] Yin, H., Wang, B., & Song, J. (2024). Prioritized experience in multi-agent DSA using RL. *IEEE Transactions on Wireless Communications*, 23(3), 2105–2117.
<https://doi.org/10.1109/TWC.2024.3290105>
- [20] Sun, X., Zhang, T., & Han, Y. (2025). Graph reinforcement learning for RIS-assisted intelligent beam selection. *IEEE Transactions on Communications*, 73(2), 1032–1045.
<https://doi.org/10.1109/TCOMM.2025.3340178>